

Bringing Genetics and Genomics to Crop Simulations: Experiences with Wheat, Sorghum and Common Bean in Solving the GEM-to-P Problem

J. W. White

(US Arid Land Agricultural Research Center, USDA-ARS, 21881 North Cardon Lane,
Maricopa, AZ 85239, USA E-mail: jeffrey.white@ars.usda.gov)

Abstract: Genetics and genomics offer avenues to reduce model uncertainty by improving descriptions of cultivar differences and of individual plant processes. Ultimately, this should greatly enhance our ability to integrate data on genotypes (G), environment (E) and management (M) in order to quantitatively predict phenotypes (P), which can be termed the GEM-to-P problem. This paper reviews use of genetics and genomics with emphasis on wheat (*Triticum aestivum*), sorghum (*Sorghum vulgare*) and common bean (*Phaseolus vulgaris*). Cultivar specific parameters, such as for photoperiod sensitivity or grain size, are often problematic because their values are determined empirically from field studies and because the assumed physiology is inaccurate. Genotypic data should be more reliable than phenotypic data, since environmental effects are minimized. Using the Cropping System Model (CSM) for bread wheat, sorghum and common bean, coefficients were estimated through linear functions of gene effects. For all three crops, simulations using gene-based coefficients were similar to those from conventional coefficients. The main constraint to wider use of this approach is the limited number of loci that have been characterized for readily modeled traits. However, emerging genomic tools allow rapid, robust characterization of genes such as the *Vrn* and *Ppd* series in wheat, and data limitations are diminishing. Genomics can also improve how processes are modeled. Examples include determining the end of the juvenile phase, characterizing interactive effects of temperature on photoperiod sensitivity, improving how tiller development is modeled, and estimating carbon costs of low-lignin traits for bioenergy crops. Realizing the potential of crop genetics and genomics, however, will not happen spontaneously. Modelers must broaden their understanding of genomics and related fields, while developing effective collaborations with the plant biology community.

Keywords: QTL, quantitative trait loci

1 Introduction

Process-based ecophysiological models are powerful tools for predicting how plant performance (the phenotype) vary in response to genetic makeup of the plants (G), the biophysical environment (E) they grow in, and how the plants are managed (M). Cooper et al. (2002) described this prediction problem as the genotype to phenotype or GP problem, but it is more accurately termed the GEM-to-P problem, recognizing the contributions of environment and management. While we are far from solving the GEM-to-P problem, ecophysiological models are widely used both as research and decision support tools.

Uncertainty erodes confidence in crop simulations but can be decreased by accurately specifying inputs and improving how individual processes are represented. Genetics and genomics offer

avenues to benefit both aspects of modeling (White and Hoogenboom, 2003). The simplest and most immediately practicable opportunities relate to improved estimation of model parameters used to represent cultivar differences, such as for photoperiod sensitivity or grain size. These parameters are often problematic because their values are determined empirically from field studies and because the assumed underlying physiology is inaccurate. Further opportunities reside in using information from genomics and related fields to improve representations of specific physiological processes.

This paper reviews use of genetics and genomics to improve models with emphasis on experiences from bread wheat (*Triticum aestivum*), sorghum (*Sorghum vulgare*), and common bean (*Phaseolus vulgaris*). Table 1 describes six levels of genetic detail that models may implement, ranging from non-species specific, generic models to models that seek to describe processes by scaling up from gene-sequences or gene products and eventually, to major physiological processes or even an entire plant. These six levels provide a useful framework for our discussion. Current widely-used models such as the Cropping Systems Model (Jones et al., 2003), which is distributed with the Decision Support System for Agrotechnology Transfer for (DSSAT) software (Hoogenboom et al., 2004), and APSIM (McCowan et al., 1996) are at level 3 (Table 1), where genetic differences are specified through empirically adjusted coefficients affecting traits like photoperiod sensitivity, leaf appearance rate, and characteristic leaf or grain sizes. Emphasis is given here to improvements in modeling per se. The ultimate goal, of course, is to benefit applications ranging from ideotype specification to cultivar adaptation to climate change (Hoogenboom et al., 2004).

Table 1 Six levels of genetic detail in crop simulations models (after White and Hoogenboom, 2003)

-
- | |
|--|
| 1. Generic: no consideration of difference among species |
| 2. Species-specific: species described but no cultivar differences recognized |
| 3. Cultivar specific parameters used to describe cultivar differences |
| 4. Cultivar differences represented by genotypes with linear model parameters |
| 5. Cultivar differences expressed through processes described using knowledge of gene expression and gene products |
| 6. Full representation of gene regulators, gene-products, etc. in networks |
-

2 Gene-Based Estimation of Model Parameters

The most immediately practicable approach for employing genetic information is to estimate cultivar parameters as a function of the alleles present at known loci, corresponding to level 4 in Table 1. An essential first step is to assemble field data for a large set of germplasm known to vary for loci of interest. This can be problematic since breeding nurseries typically deal with materials of similar adaptation and may show limited variation. In the case of sorghum, near-isogenic lines for maturity loci had been tested in a wide range of environments, providing a valuable data resource for model development (White et al., 2007b). The field data are then used to calibrate the cultivar coefficients using the conventional approach of adjustment and comparison of measured vs. simulated data on phenotypes. With the coefficient values in hand, one can then proceed to estimate the effects of known loci.

In the simplest case, where only dominant and recessive alleles are known and cultivars are pure, inbred lines, the alleles can be scored with a value of 1 if dominant and 0 if recessive, and their effects estimated through linear regression. Thus, for two alleles affecting a cultivar parameter P , one might have:

$$P = a + bL_1 + cL_2$$

where a , b and c are regression coefficients (estimated through ordinary least squares), and L_1 and L_2 are scores for two loci. Interactions among loci (genetic epistasis) are incorporated by multiplying loci effects (e.g., $L_1 \times L_2$), as done for the *Ppd* and *Hr* loci affecting photoperiod sensitivity in common bean (White and Hoogenboom, 1996). Applied to wheat, the estimating equation for the effect of the *Ppd-D1* locus on the photoperiod sensitivity coefficient PID thus was:

$$PID = 39.9 - 13.9 Ppd-D1 \quad R^2 = 0.42^{**}$$

where *Ppd-D1* indicated the effect of the corresponding locus. Note that dominance (*Ppd-D1* = 1) reduces the value of PID, reducing photoperiod sensitivity, which agrees with the described effect of the locus.

Selection of alleles used to estimate a parameter should emphasize physiological rationales. For example, based on our knowledge of action of the *Rht* loci in wheat, one might expect them to influence leaf size but not photoperiod sensitivity. Tests for significance of coefficients can be used to reject loci whose effects are too small or uncertain to merit inclusion in a given regression, but physiological understanding (or at least firm hypotheses) should lead variable (locus) selection.

Estimation through linear regression has been used successfully for common bean, soybean (*Glycine max*), sorghum and wheat (White and Hoogenboom, 1996; Hoogenboom et al., 1997; Stewart et al., 2003; Messina et al., 2006; White et al., 2007b; White et al., 2008). Fig. 1 shows simulations of days to flowering using gene-based approaches for sorghum and wheat, using independent evaluation datasets. For both crops, simulations of flowering dates were similar to those obtained with conventional cultivar coefficients (White et al., 2007b and 2008).

Table 2 Comparisons of results for conventional and gene-based simulation models of time to flower for sorghum (White et al., 2007b), and wheat (White et al., 2008). Evaluation datasets are independent from calibration sets, predominantly involving different locations. The variation that was explained by assuming no differences among cultivars was tested by evaluating simulations based on a single generic cultivar

Crop	Dataset	N	Model type	Mean	R ²	Slope	RMSE
Sorghum	Calibration	108	Generic	64	0.23**	0.76	12.5
			Conventional	67	0.61**	0.76	8.8
			Gene-based	67	0.55**	0.73	9.5
	Evaluation	74	Generic	58	0.31**	1.41	8.8
			Conventional	61	0.60**	0.91	6.7
			Gene-based	61	0.61**	0.97	6.6
Wheat	Calibration	540	Generic	208	0.95**	0.95	9.7
			Conventional	209	0.98**	0.95	6.6
			Gene-based	208	0.96**	0.92	8.6
	Evaluation	1499	Generic	214	0.89**	1.02	10.4
			Conventional	213	0.92**	1.00	9.0
			Gene-based	214	0.90**	0.99	9.9

** Significant at the $P < 0.01$ level.

Table 2 summarizes results of simulations of time to flowering in sorghum and wheat using conventional and gene-based modeling. In testing effects of genotypes, the effect of a "generic" cultivar was included in order to provide a more meaningful basis for comparison. For each species,

parameter values for the generic cultivar were estimated as the mean values for all cultivars being examined (the modal value could also be used). For sorghum, the generic cultivar performed poorly compared to the conventional and gene-based models, but in wheat, the generic cultivar was surprisingly effective (Table 2). One interpretation of these results is that for the wheat dataset, there were very large differences among environments that a generic cultivar was capable of representing, whereas in sorghum cultivar differences were large relative to environmental differences. Regression analysis and related techniques provide valuable tool for exploring the relative ability of models to explain genotype, environment and management effects (White et al., 2007a).

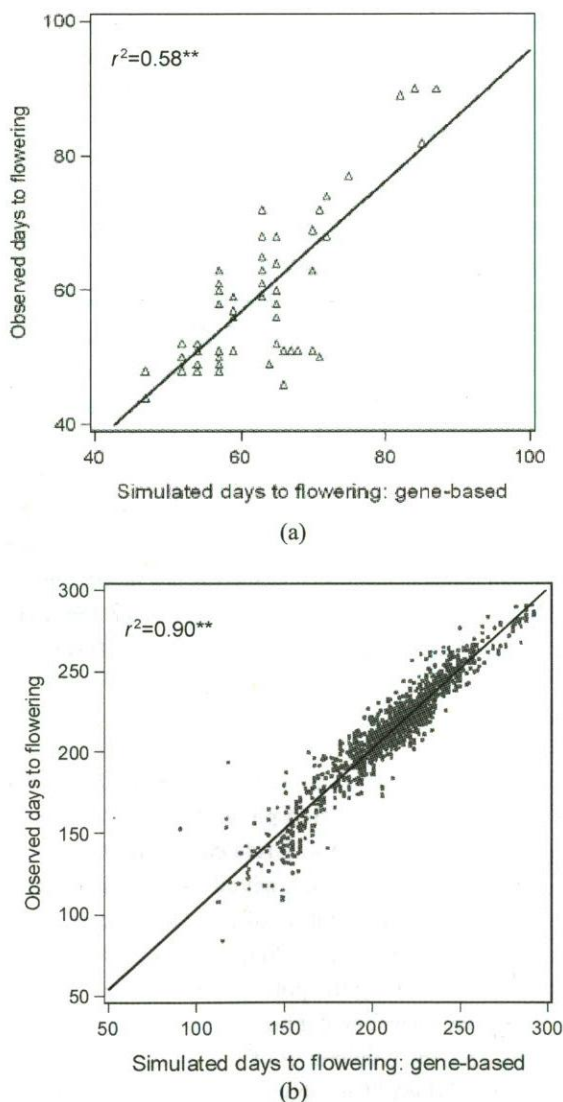


Figure 1 Comparison of observed vs simulated days to flowering using gene-based model parameters. Both datasets were independent of calibration data. (a) For sorghum recombinant inbred lines varying for *Ma* loci (White et al., 2007b). (b) For 30 wheat cultivars of the IWWPN (White et al., 2008)

In GeneGro, the common bean model that first used gene-based estimation, genotypes were input directly into the model, and gene effects were specified as part of the model code. Subsequently, we realized that since the gene effects were only used to specify model parameters, the calculations

should be external to the model code: the gene-based estimates of cultivar parameters can simply replace the conventionally estimated parameters. Thus for wheat and sorghum, we estimated gene-effects externally and inputted the values to the model through alternate files containing gene-based estimates of the cultivar parameters. Additional minor innovations in the basic approach have included consolidating effects of three loci with similar effects (*Vrn-A1*, *Vrn-B1*, and *Vrn-D1* in wheat) through a single variable for number of dominant alleles present (White et al., 2008) and accounting for the three alleles of the sorghum *Ma3* locus by scoring values of 1.0, 0.5 and 0 based on estimates of the relative effect of each allele (White et al., 2007b).

The main constraints to gene-based estimation of cultivar parameters probably are that many modelers are not familiar enough with crop genetics to implement the approaches, that very few loci are sufficiently characterized to permit efficient modeling, and that even if loci are well characterized in terms of their effects and underlying physiology, few cultivars have been genotyped for loci of interest. Reducing the first constraint requires effort from the modeling community. The latter two constraints are diminishing rapidly as genomic tools facilitate characterization of cultivars for loci such as the *Vrn* and *Ppd* series in wheat as well as accelerate identification of new loci. Illustrating this progress, CIMMYT is now capable of genotyping wheat lines for photoperiod sensitivity (*Ppd-B1*, *Ppd-D1*), vernalization (*Vrn-A1*, *Vrn-B1*, *Vrn-D1*) and plant height (*Rht-B1*, *Rht-D1*), at a cost per locus of \$2.00 US (S. Dreisigacker, 2008, personal communication).

In this transition period when characterization of loci affecting physiological traits limits progress in modeling, a complementary strategy is to use molecular markers to identify quantitative trait loci (QTL), and then use loci effects to estimate model parameters. In GeneGro (White and Hoogenboom, 1996), this approach was implicitly used since the three hypothetical loci for seed mass were inferred mainly from QTL studies. QTL-based estimation has been used in simulating barley growth and yield (*Hordeum vulgare*; Yin et al., 2000), sorghum phenology (Chanterreau et al., 2001) and leaf growth in maize (Reymond et al., 2003). Traditional QTL studies based on populations of inbred lines have numerous limitations, but these problems are being overcome through approaches combining high throughput marker systems such as Diversity Array Technology (Wenzl et al., 2004) with association mapping (Janink and Walsh, 2002; Crossa et al., 2007).

3 Improved Physiology

In a sense, simple linear models of gene action imply partial knowledge of the underlying physiology. Specifying that the bean *Ppd* locus only affected photoperiod parameters implicitly assumed that the photoperiod response is independent of traits related to growth such as intrinsic leaf size or potential leaf photosynthetic rate. However, such trivial examples represent minimal advances in our understanding of physiology. Better integration of plant science and ecophysiological modeling is still needed (Hammer et al., 2004; Yin et al., 2004; White, 2006) in order to move to level 5 of Table 1. Four cases are explored illustrating the potential benefits to modeling.

Onset of flowering in many plant species depends on activation of the meristem identity gene *APETALA1* (*API*), which initiates the conversion of the shoot apex to a reproductive state (Murai et al., 2003; Imaizumi and Kay, 2006). The photoperiod system inhibits *API*, but to model the photoperiod effect, one needs to specify time of onset and end of photoperiod sensitivity. The onset of sensitivity is considered to represent the end of the juvenile phase and usually is determined with experiment manipulations. Plants of a sensitive genotype are grown under both short and long day regimes, and individual plants are transferred from one regime to the other every few days. Analysis of days to flowering (or other stages) should reveal a date prior to which the plants can be grown in either photoperiod with no effect on development, indicating that this date corresponds to the end of the juvenile phase (e.g., Alagarswamy et al., 1998). Besides being labor intensive and requiring

large numbers of plants, the experiments may not reveal a clear transition (see Figs. 2 and 3 (in Alagarswamy et al., 1998)). Genomic tools should allow tracking the activity of the photoperiod system as manifested in levels of specific gene-products and ultimately provide reliable criteria for measuring the end of the juvenile phase, as well as the end of photoperiod sensitivity much later in development.

Warmer temperatures increase photoperiod sensitivity in many crops, and there are genetic differences among cultivars for this effect (e.g., Quinby et al., 1973; Cao and Moss, 1989; White et al., 1996). Presumably due to uncertainty over its physiological description, however, few models consider this interaction. Genomics offers the possibility of clarifying the underlying response mechanisms. Simply improving characterization of genotypes would facilitate testing alternative approaches to model the interaction, but genomics might determine whether the response involves a specific, independent temperature sensor or simply results from generic effects of temperature on metabolic rates.

Modeling effects of tillers is also problematic. Many models work well for high population densities and narrow row spacings, but they underestimate the growth and yield of tillers when interplant competition is low. The *Tin* locus inhibits tillering in wheat (Richards, 1988) and provides one entry point for molecular-level dissection of tillering. Tiller production involves phytochrome B as part of the shade avoidance syndrome, and advances in understanding of this system (Franklin and Whitelam, 2005) are another source of information for modeling tiller development.

A fourth example involves estimating the carbon costs of novel traits. Changes in plant composition are actively sought as a means to improve the digestibility of cereal straw for cellulosic ethanol production (Sarath et al., 2008). Lignins, as complex polymers of phenolic compounds, are not fermentable, and breeding seeks to manipulate their levels by means such as the brown midrib mutants of maize and sorghum (Sarath et al., 2008). In cereals, lignins typically represent 15 to 20% of stem dry weight, so reducing their levels should reduce the net growth respiration. The pioneering work of F. Penning de Vries suggested that the cost of biosynthesis of lignins is approximately 2.2 g glucose per 1 g lignin vs. 1.2 g g⁻¹ for cellulose (Penning de Vries et al., 1983). These values allow a rough estimation of the impact of reduced lignin concentration on growth respiration, but more precise estimates would require a detailed understanding of the metabolic pathways affected—information obtainable through genomics and metabolomics. Of course, the metabolic benefit of lowering lignin concentration should be considered in the light of possible adverse effects such as reduced stem strength. Note that consideration of costs of biosynthesis should have broad applications in assessing possible trade-offs for crop improvement efforts that seek to radically alter plant structure or composition, regardless of whether the changes are through application of molecular or conventional breeding.

4 Plant Systems Biology

The sixth and highest level of complexity in modeling approaches is to explicitly represent biochemical networks of genes and gene products to achieve what is sometimes termed the “in silico plant” (Minorsky, 2003). Various software systems can simulate complex metabolic pathways (e.g., Hoops et al., 2006; Takahashi et al., 2004), and the Systems Biology Markup Language (SBML) was established to facilitate this research (Hu et al., 2005). Welch et al. (2003 and 2005) have probably advanced the furthest in applications that are relevant to crop modeling through their efforts to model control of flowering in *Arabidopsis* with simplified networks, although other flowering models also consider biochemical pathways (e.g., Locke et al., 2005).

5 How Well Can We Solve the GEM-to-P Problem?

While the arguments outlined above support optimism for improved ability to predict phenotypes from genotypes, environment and management, there is value in considering the potential limits to model prediction, especially using genetic data as inputs. For a given quantitative trait, one might ask what percent of variation can ultimately be explained by models that integrate the best available knowledge about genetics, physiology and the environment (Fig. 2). Certainly there has been progress since Reaumur (1735) first proposed concepts that provided the foundation of quantitative modeling of plant development, but if we look to 2050 or beyond, how much further improvement can we expect?

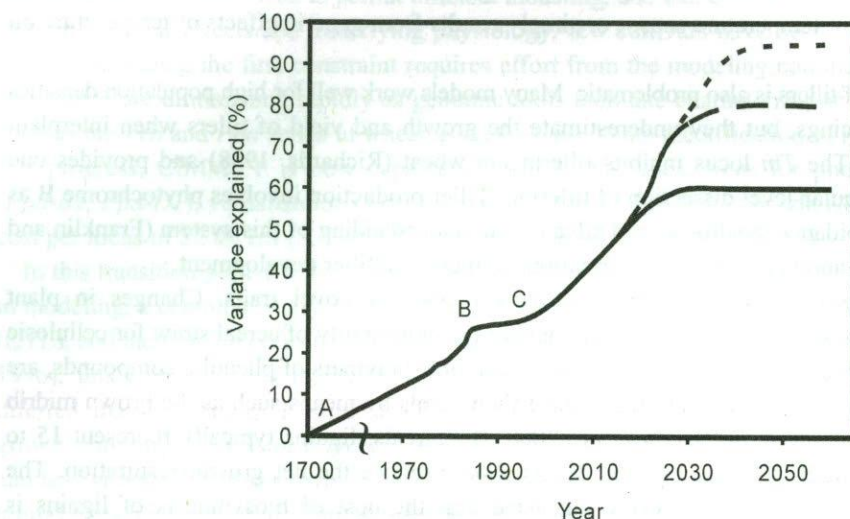


Figure 2 Hypothetical plot of progress in simulation of an arbitrary trait (e.g., flowering date). Note interrupted axis between 1700 and 1970. The three points correspond to the following events: A. First conceptualization of accumulated temperature affecting phenology by Reaumur (1735). B. Expansion of simple developmental models to consider photoperiod. C. First attempts to introduce genetic information into models

When the *Arabidopsis* genome was sequenced in 2000 (The *Arabidopsis* Genome Initiative, 2000), the genome was estimated to contain 25,000 genes. Subsequent progress in genomics suggests that the notion of “gene” is more complex than expected. MicroRNAs have proven to have important signaling roles, complicating the “one gene, one enzyme” paradigm (Axtell et al., 2007). Similarly, alternate splicing patterns of pre-messenger RNA mean that a single gene sequence can result in multiple transcripts and hence, multiple proteins (Reddy, 2007). And demonstration of important epigenetic effects imply that gene sequences are sometimes less important in regulating transcription than how DNA is physically packaged (Bender, 2002).

Taking 25,000 genes as a lower estimate of the potential complexity of a single species, one might argue that the prospects for major improvements in predictive ability are poor. However, counter-arguments in favor of additional simplifications exist. Many gene-systems may primarily affect disease or pest resistance and have minimal direct relevance to processes considered in ecophysiological models. Additionally, the genetic systems for pollination, fertilization and early embryo development may have limited relevance except as they relate to specific stress responses. In many cases, signaling systems, such as from roots to shoots, may not require complete description: it should suffice to know that a sensor exists and that a message is transmitted to a receptor. The simplified gene network model of Welch et al. (2005) for phenology exemplifies this “middle way.”

The possible limits of crop modeling to solve the GEM-to-P problem also relates to the broader issue of whether evolution is driven by slow accumulation of micro-adaptations, which would favor arguments for genetic complexity, or whether most mutations are essentially neutral, and only occasionally do mutations cause major, step-like changes in phenotypes. The first hypothesis implies that we must ultimately model massively complex regulatory networks, while the latter favors a view that major simplifications are possible. The increasing availability of genome sequence data allows more rigorous testing of the nature of evolution (e.g., *Drosophila* 12 Genomes Consortium, 2007), including for crop domestication (Wright et al., 2005). To date, the evidence mainly suggests that the genomic rate of adaptive evolution differs greatly among species (Eyre-Walker, 2006).

Prospects for bringing genetics and genomics to ecophysiological modeling, and solving the GEM-to-P problem, appear excellent. However, the required research is surrounded with uncertainty that may ultimately require solving fundamental issues of plant evolution and function. Furthermore, realizing this potential will not happen spontaneously. Modelers must broaden their understanding of genomics and related fields, while developing effective collaborations with the plant biology community.

References

- Alagarswamy, G., Reddy D.M., and Swaminathan G. (1998) Durations of the photoperiod-sensitive and -insensitive phases of time to panicle initiation in sorghum. *Field Crops Res.* 55:1 – 10.
- Arabidopsis Genome Initiative. (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* (London) 408:796 – 815.
- Axtell, M.J., Snyder J.A., and Bartel. D.P. (2007) Common functions for diverse small RNAs of land plants. *Plant Cell* 19:1750 – 1769.
- Bender, J. (2002) Plant epigenetics. *Current Biol.* 12:R412-R414.
- Chanterneau, J., Trouche G., Rami J.F., et al. (2001) RFLP mapping of QTLs for photoperiod response in tropical sorghum. *Euphytica* 120:183 – 194.
- Cooper, M., Chapman S.C., Podlich D.W. et al. (2002) The GP problem: quantifying gene-to-phenotype relationships. *In Silico Biol.* 2:151 – 164.
- Crossa, J., Burgueno J., Dreisigacker S. et al. (2007) Association analysis of historical bread wheat germplasm using additive genetic covariance of relatives and population structure. *Genetics* 107:1889 – 1913.
- Drosophila* 12 Genomes Consortium. (2007) Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* 450:203 – 218.
- Franklin, K.A., and Whitelam. G.C. (2005) Phytochromes and shade-avoidance responses in plants. *Ann. Bot.* 96:169 – 175.
- Hammer, G.L., Sinclair, T.R., Chapman, S.C., et al. (2004) On systems thinking, systems biology, and the in silico plant. *Plant Phys.* 134:909 – 911.
- Hoogenboom, G., Jones J.W., Wilkens P.W., et al. (2004b) Decision Support System for Agrotechnology Transfer Version 4.0 [CD-ROM]. University of Hawaii, Honolulu, HI.
- Hoogenboom, G., and White. J.W. (2003) Improving physiological assumptions of simulation models by using gene-based approaches. *Agron. J.* 95:82 – 89.
- Hoogenboom G., White J.W., Acosta-Gallegos J., et al. (1997) Evaluation of a crop simulation model that incorporates gene action. *Agron. J.* 89:613 – 620.
- Hoogenboom, G., White J.W., and Messina C.D. (2004a) From genome to crop: integration through simulation modeling. *Field Crops Res.* 90:145 – 163.
- Hoops, S., Sahle S., Gauges R., et al. (2006) COPASI- A COMplex PATHway Simulator. *Bioinformatics* 22: 3067 – 3074.
- Hucka, M., Finney A., Sauro H.M., et al. (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* 19:524 – 531.
- Imaizumi, T., and Kay S.A. (2006) Photoperiodic control of flowering: not only by coincidence. *Trends Plant Sci.* 11:550 – 558.

- Jannink, J.-L. and Walsh, J.B. (2002) Association mapping in plant populations. In: Kang M.S. (ed.) Quantitative Genetics, Genomics and Plant Breeding, edited by CABI, Wallingford, UK.
- Jones, J.W., Hoogenboom G., Porter C.H., et al. (2003) The DSSAT Cropping System Model. *Eur. J. Agron.* 18:235 – 265.
- Locke, J.C.W., Southern M.M., Kozma-Bognar L., et al. (2005) Extension of a genetic network model by iterative experimentation and mathematical analysis. *Mol. Syst. Biol.* 1:2005.0013.
- McCowan, R.L., Hammer G.L., Hargreaves J.N.G., et al. (1996) APSIM: a novel software system for model development, model testing and simulation in agricultural systems research. *Agric. Sys.* 50:255 – 271.
- Messina, C.D., Jones J.W., Boote K.J., et al. (2006) A gene-based model to simulate soybean development and yield responses to environment. *Crop Sci.* 46:456 – 466.
- Minorsky, P.V., (2003) Achieving the in silico plant. *Systems biology and the future of plant biological research. Plant Physiol.* 132, 404 – 409.
- Quinby, J.D., Hesketh J.D., and Voigt, R.L. (1973) Influence of temperature and photoperiod on floral initiation and leaf number in sorghum. *Crop Sci.* 13:243 – 246.
- Reaumur, R.A.F.d. (1735) Observations du thermometre, faites a Paris pendant l'annee 1735, compares avec celles qui ont ete faites sous la ligne, a l'Isle de France, a Alger et en quelques-unes de nos isles de l'Amerique. *Mem. Acad. des Sci., Paris* 545 – 576.
- Reddy, A.S.N. (2007) Alternative splicing of pre-messenger RNAs in plants in the Genomic Era. *Annual Rev. Plant Biol.* 58:267 – 294.
- Reymond, M., Muller B., Leonardi A., et al. (2003) Combining quantitative trait loci analysis and an ecophysiological model to analyze the genetic variability of the responses of maize leaf growth to temperature and water deficit. *Plant Physiol.* 131:664 – 675.
- Richards, R. (1988) A tiller inhibitor gene in wheat and its effect on plant growth. *Austral. J. Agric. Res.* 39:749 – 757
- Sarath, G., Mitchell R., Sattler S., et al. Opportunities and roadblocks in utilizing forages and small grains for liquid fuels. *J. Industr. Microbiol. Biotech.* 10.1007/s10295-007-0296-3.
- Stewart, D.W., Cober E.R., and Bernard R.L. (2003) Modeling genetic effects on the photothermal response of soybean phenological development. *Agron J* 95:65 – 70.
- Takahashi, K., Kaizu K., Hu B., et al. (2004) A multi-algorithm, multi-timescale method for cell simulation. *Bioinformatics* 20:538 – 546.
- Welch, S.M., Roe J.L., and Dong Z. (2003) A genetic neural network model of flowering time control in *Arabidopsis thaliana*. *Agron. J.* 95:71 – 81.
- Welch, S.M., Roe J.L., Das S., et al. (2005) Merging genomic control networks and soil-plant-atmosphere-continuum models. *Agric. Syst.* 86:243 – 274.
- Wenzl, P., Carling J., Kudrna D., et al. (2004) Diversity Arrays Technology (DArT) for whole-genome profiling of barley. *PNAS* 101:9915 – 9920.
- White, J.W., Kornegay J., and Cajiao C. (1996) Inheritance of temperature sensitivity of the photoperiod response in common bean (*Phaseolus vulgaris*). *Euphytica* 91:5 – 8.
- White, J.W. (2006) From genome to wheat: Emerging opportunities for modeling wheat growth and development. *Eur. J. Agron.* 25:79 – 88.
- White, J.W., and Hoogenboom. G. (1996) Simulating effects of genes for physiological traits in a process-oriented crop model. *Agron. J.* 88:416 – 422.
- White, J.W., and Hoogenboom. G. (2003) Gene-based approaches to crop simulation: past experiences and future opportunities. *Agron. J.* 95:52 – 64.
- White, J.W., Boote K.J., Hoogenboom G., et al. (2007a) Regression-based evaluation of ecophysiological models. *Agron. J.* 99:419 – 427.
- White, J.W., Hoogenboom G., and Ottman M. (2007b) Modeling phenology of sorghum based on known maturity (*Ma*) loci. P. 83-85. In: Farming Systems Design 2007. Proc. Int. Conf., Catania, Italy. 10-12 Sept. 2007. http://www.iemss.org/farmsys07/uploads/Main/Field_farm_scale_CD.pdf (verified 31 March 2008).

- White, J.W., Herndl M., Hunt L.A., et al.(2008) Simulation-based Analysis of Effects of *Vrn* and *Ppd* Loci on Flowering in Wheat. *Crop Sci.* 48: 678 – 687.
- Yin, X., Chasalow S.D, Dourleijn C.J, et al.(2000) Coupling estimated effects of QTLs for physiological traits to a crop growth model: predicting yield variation among recombinant inbred lines in barley. *Hered.* 85:539 – 549.
- Yin, X., Struik P.C, and Kropff M.J. (2004) Role of crop physiology in predicting gene-to-phenotype relationships. *Trends in Plant Sci.* 9:426 – 432.